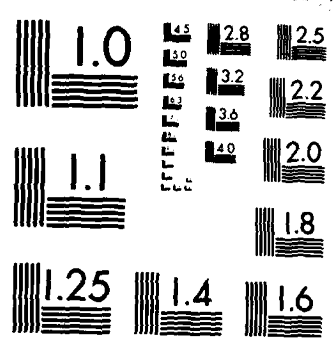
AD-A153 534 LEAST-SQUARES APPROXIMATION TO MINIMUM CHI-SQUARE ESTIMATORS OF LOCATION A..(U) WISCONSIN UNIV-MADISON MATHEMATICS RESEARCH CENTER A E MUHLY ET AL. FEB 85 1/1

UNCLASSIFIED MRC-TSR-2792 DAAG29-80-C-0041 F/G 12/1 NL

END

FILMED

DTIC


1.0
2.8
2.5
3.2
2.2
3.6
4.0
2.0
1.1
1.8
1.25
1.4
1.6

(2)

AD-A153 534

MRC Technical Summary Report #2792

LEAST-SQUARES APPROXIMATION TO MINIMUM CHI-SQUARE ESTIMATORS OF LOCATION AND SCALE PARAMETERS AND THEIR EFFECT ON THE PEARSON CHI-SQUARE TEST

A. E. Muhly and J. Gurland

Mathematics Research Center
University of Wisconsin—Madison
610 Walnut Street
Madison, Wisconsin 53705

February 1985

(Received December 6, 1983)

Sponsored by

U. S. Army Research Office
P. O. Box 12211
Research Triangle Park
North Carolina 27709

UNIVERSITY OF WISCONSIN-MADISON
MATHEMATICS RESEARCH CENTER

LEAST-SQUARES APPROXIMATION TO MINIMUM CHI-SQUARE ESTIMATORS OF LOCATION AND SCALE PARAMETERS AND THEIR EFFECT ON THE PEARSON CHI-SQUARE TEST

A. E. Muhly and J. Gurland

Technical Summary Report #2792
February 1985

ABSTRACT

Application of the Pearson chi-square test to goodness of fit of a distribution often leads to serious difficulties, particularly in the formation of intervals (as in the case of a continuous distribution) and in the estimation of unknown parameters. Under suitable conditions and with appropriately constructed estimators of the parameters, the test statistic converges in distribution to that of chi-square as the sample size increases. In the present paper, a comparatively simple least-squares approximation to the minimum chi-square estimator is developed which, when appropriately implemented, results in an asymptotic chi-square distribution of the test statistic. This estimator is developed for the cases of fixed and random intervals, and the role of the underlying assumptions is studied in detail.

AMS (MOS) Subject Classifications: 62E20, 62F12, 62F03

Key Words: Pearson chi-square test, test of fit, asymptotic distribution, least-squares approximation.

Work Unit Number 4 (Statistics and Probability)

A-1

Sponsored by the United States Army under Contract No. DAAG29-80-C-0041.

SIGNIFICANCE AND EXPLANATION

The research considered here falls within the area of statistics known as tests of goodness of fit. Since many widely used statistical procedures are based on particular distributional assumptions, it is quite possible in practice that the resulting analysis can lead to erroneous or misleading conclusions. In virtually every area where statistics is applied, e.g. reliability of a product, prediction of performance, evaluation of a treatment, etc., some sorts of assumptions are usually made, either explicitly or implicitly.

One of the most widely used tests of goodness of fit is that known as the Pearson chi-square test. Many articles on this subject have appeared in the statistical literature, dealing with such problems as the construction and number of intervals to be used in the test, and the effect of estimators of unknown parameters on its asymptotic distribution. In the present paper, a comparatively simple estimator, based on least-squares, is developed which, under suitable conditions precisely formulated, result in an asymptotic distribution of the test statistic which is that of chi-square. The usefulness of this result is in improving the application of the Pearson chi-square test by utilizing the simpler asymptotic chi-square distribution which obviates the difficulty of trying to employ a test statistic whose asymptotic distribution depends on unknown parameters.

The responsibility for the wording and views expressed in this descriptive summary lies with MRC, and not with the authors of this report.

LEAST-SQUARES APPROXIMATION TO MINIMUM CHI-SQUARE ESTIMATORS OF LOCATION AND SCALE PARAMETERS AND THEIR EFFECT ON THE PEARSON CHI-SQUARE TEST

A. E. Muhly and John Gurland

1. Introduction

Suppose that a random sample of size N from a population with unknown distribution function $F(\cdot)$ is given and that it is desired to test the composite null hypothesis

$$H_o : F(\cdot) = F_o(\cdot;\theta). \tag{1.1}$$

In (1.1) it is assumed that the functional form of F_o is completely specified while the parameter vector θ, which is an element of an open set Θ contained in s-dimensional Euclidean space, is unknown and unspecified. In addition, it is assumed that for every θ in Θ, $F_o(x;\theta)$ is continuous in x.

To apply the Pearson chi-square test to the above problem it is necessary to: i) partition the support of F_o into k intervals, and ii) estimate the unknown value of θ. When the endpoints of the intervals are specified prior to observing the sample (the fixed interval case) it is well known (see Fisher (1924), Cramér (1946), Rao (1958), and Birch (1964)) that if the minimum chi-square (or asymptotically equivalent) estimator of θ is employed then the limit distribution of the resulting test statistic is the chi-square distribution with $k-s-1$ degrees of freedom. Furthermore, Chibisov (1971), pursuing a conjecture of Watson's (1958, 1959), showed that this result remains true when the interval endpoints are allowed to depend on the sample (the random interval case), provided that these endpoints are treated as fixed for the purpose of obtaining this estimator.

The practical significance of the above results is diminished, however, since it is rarely possible to obtain closed form expressions for the minimum chi-square estimator. This fact led to a discussion of the effect that other methods of estimation have on the

---

Sponsored by the United States Army under Contract No. DAAG29-80-C-0041.

limit distribution of the Pearson chi-square test statistic. In the fixed interval case the definitive result along these lines is due to Chernoff and Lehmann (1954) who showed that if the method of maximum likelihood based on the original (i.e. ungrouped) sample is employed to estimate θ then the desired limit distribution is the same as the distribution of

$$\chi^2_{k-s-1} + \sum_{i=1}^{s} \lambda_i z_i^2 \tag{1.2}$$

where χ^2_{k-s-1} is a random variable distributed as chi-square with $k-s-1$ degrees of freedom, $Z_1,\ldots,Z_s$ are independent standard normal random variables distributed independently of χ^2_{k-s-1}, and $\lambda_1,\ldots,\lambda_s$, which are the characteristic roots of a certain matrix, satisfy $0 < \lambda_i < 1$. Roy (1956) and, working independently, Watson (1957, 1958) showed that this result also holds true in the random interval case with one important distinction. While it is generally the case that the λ_i appearing in (1.2) depend on the unknown value of θ, if F_0 depends only on unknown location and scale parameters and if random interval boundaries, formed as in (3.17) below, are employed then the λ_i appearing in (1.2) do not depend on θ.

Other estimators, satisfying fairly general regularity conditions, have also been considered in the random interval case (see, in particular, Roy (op. cit.) and Moore and Spruill (1975)). The resulting limit distribution of the Pearson chi-square test statistic when these estimators are used is similar in form to the distribution of (1.2) except that now the λ_i appearing there need not be bounded above by one. Furthermore, it is generally the case that the λ_i depend on θ although, as shown by Dahiya and Gurland (1972), this is not the case when only location and scale parameters are involved and they are estimated by the ungrouped sample mean and standard deviation.

Published attempts to contend with the problems raised in the preceding paragraphs include the following. If the number, k, of intervals employed is large then it is possible to use either i) Cramér's (op. cit.) approximation to the minimum chi-square

estimator for location and scale parameters, or ii) Watson's (1958) argument for ignoring the term $\sum_{i=1}^{s} \lambda_i Z_i^2$ appearing in (1.2). If, on the other hand, k is moderate or small then, as suggested by Rao and Robson (1974) and generalized by Moore (1977), it is possible to alter the quadratic form which defines the Pearson chi-square statistic to obtain a limiting chi-square distribution. Alternatively, if it is known that (1.2) does not depend on θ, then percentiles of its distribution can be computed by means of Laguerre series.

The solutions mentioned above are not entirely adequate. First, it is demonstrated in Dahiya and Gurland (1973) that the power of the Pearson chi-square test may be drastically reduced by using a large number of intervals. This observation is further supported by Spruill (1976a) who uses a different method for comparing power. Therefore, it is not desirable to assume, a priori, that k is large. Second, in some cases of interest the elements of the matrix of the modified quadratic form used by Rao and Robson (op. cit.) and the λ_i appearing in (1.2) pose significant computational problems. Finally, it is possible that the use of the minimum chi-square estimator leads to a more powerful test than any of the solutions mentioned above (see Chibisov (op. cit.), Moore and Spruill (op. cit.), and Spruill (1976b)).

In this paper it is shown how, in some situations, it is possible to obtain a least squares approximation, $\tilde{\tilde{\theta}}_N$, to the minimum chi-square estimator $\tilde{\theta}_N$ as defined in (3.5) and (3.22) below. Initially, it will be assumed that F_o depends only on unknown location and scale parameters. It will be seen that $\tilde{\tilde{\theta}}_N$ is easy to compute and, when random intervals are employed, has a particularly nice interpretation. Furthermore, it will be shown that $\tilde{\tilde{\theta}}_N$ and $\tilde{\theta}_N$ are asymptotically equivalent in the sense that $\tilde{\tilde{\theta}}_N - \tilde{\theta}_N = o_p(1/\sqrt{N})$, and, therefore, that the limit distribution of the Pearson chi-square statistic is the chi-square distribution with k-3 degrees of freedom when $\tilde{\tilde{\theta}}_N$ is used.

The next section introduces some conventions that will be used throughtout this paper and presents the assumptions that are required. Section 3 is devoted to a description of

the notation that will be used. In addition, known results are recorded which are required in later sections. In section 4, $\tilde{\theta}_N$ is presented for the random interval case and, in section 5, the necessary modifications required to define $\tilde{\theta}_N$ in the fixed interval case are given. In section 6 the asymptotic properties of $\tilde{\theta}_N$ and the resulting test statistics are derived. The assumptions given in section 2 are extended, in section 7, to treat some distributions which involve unknown parameters not of the location and scale variety. Finally, an appendix is included where certain technical points are derived.

2. Conventions and Assumptions

In this paper the following conventions will be employed. First, all vectors appearing here are column vectors and the notation A' will be used to denote the transpose of the vector (or matrix) A. In addition, if A is a vector then $\|A\| = (A'A)^{1/2}$. Second, if $a_o, a_1, \ldots, a_{k-1}, a_k$ are given numbers, the notation $\Delta a_i = a_i - a_{i-1}$ for $i = 1, \ldots, k$ will be used. If, however, a_o a_k are not defined or do not enter into the discussion, they will be treated as zero and the notation $\bar{\Delta} a_i = a_i - a_{i-1}$ for $i = 1, \ldots, k$ will be used. Thus, $\bar{\Delta} a_1 = a_1$, $\bar{\Delta} a_k = -a_{k-1}$, and $\bar{\Delta} a_i = a_i - a_{i-1}$ for $i = 2, \ldots, k-1$. Third, if X_N is a sequence of random arrays such that the elements of $f(N)X_N$ converges in probability to 0 as N tends to infinity, the notation $X_N = o_p(1/f(N))$ will be employed. Similarly, if X_N is a sequence of random arrays and if each element of $f(N)X_N$ remains bounded in probability as N tends to infinity, then the notation $X_N = o_p(1/f(N))$ will be used. Finally, if the p-dimensional random vector X has the multivariate normal distribution with mean vector μ and covariance matrix Σ, this will be denoted by $X \sim N_p(\mu, \Sigma)$. In addition, if X_N is a sequence of random vectors converging in distribution to X (where $X \sim N_p(\mu, \Sigma)$) as N tends to infinity, then this will be denoted by $X_N \xrightarrow{d} N_p(\mu, \Sigma)$. Similarly, if the random variable X has a chi-square distribution with r degrees of freedom the notation $X \sim \chi_r^2$ will be employed, while $X_N \xrightarrow{d} \chi_r^2$ means that the sequence of

random variables, X_N, converges in distribution to a random variable X (with $X \sim \chi_r^2$) as N tends to infinity.

The results obtained in sections 3 through 6 are based on the following assumptions.

<u>Assumption 1</u>: There exists a parameterization of $F_o(x;\theta)$ such that

i) $\Theta = \{\theta : \theta_1 \in \mathbf{R}$ and $\theta_2 > 0\}$

ii) $F_o(x;\theta) = F_o(\theta_1+\theta_2 x)$ where $F_o(\cdot) = F_o(\cdot;\theta)$ for $\theta = (0,1)'$.

<u>Assumption 2</u>: $F_o(x)$ is continuously differentiable in x and $f_o(x) = dF_o(x)/dx$ is positive for all finite x.

In addition to these assumptions on F_o, it will be necessary to show that various estimators of θ satisfy the following condition.

<u>Definition</u>: An estimator θ_N^{**} of θ will be said to satisfy condition A if there exists a function $h : R \to \mathbf{R}^2$ such that

i) $\int_{\mathbf{R}} h(x)dF_o(x;\theta^o) = 0$ and $\int_{\mathbf{R}} h(x)h(x)'dF_o(x;\theta^o) = G$ where the elements of G are finite and G is positive semi-definite.

ii) $\theta_N^{**} - \theta^o = \frac{1}{N}\sum_{\alpha=1}^{N} h(x_\alpha) = o_p(1/\sqrt{N})$.

Here, θ^o denotes the true unknown value of θ and $x_1,\ldots,x_N$ denotes the original sample.

The following comments are in order. i) Since the problem of primary interest is the test of H_o and not the estimation of θ, how F_o is parameterized is not a significant consideration. Thus, for example, the Cauchy distribution, with the density $(\lambda/\pi)[\lambda^2+(x-\eta)^2]^{-1}$, satisfies assumptions 1 and 2 with $F_o(x) = \frac{1}{2} + \frac{1}{\pi}\tan^{-1}(x)$, $\theta_1 = -\eta/\lambda$, and $\theta_2 = 1/\lambda$. ii) In both the fixed and random interval cases, in order to compute $\bar{\theta}_N$, a preliminary estimate, θ_N^*, of θ^o is required which satisfies $\theta_N^* - \theta^o = o_p(1/\sqrt{N})$. If $\int_{-\infty}^{\infty} x^4 d\, F_o(x;\theta^o)$ is finite, then, under the parameterization imposed by assumption 1, a natural choice for θ_N^* is given by

$$\theta_{1,N}^* = -\bar{X}_N/\sqrt{S_N^2} \quad \text{and} \quad \theta_{2,N}^* = 1/\sqrt{S_N^2} \tag{2.1}$$

where $\bar{X}_N = \frac{1}{N}\sum_{\alpha=1}^{N} X_\alpha$ and $S_N^2 = \frac{1}{N}\sum_{\alpha=1}^{N}(X_\alpha-\bar{X}_N)^2$. However, in general it is not required

that $\int_{-\infty}^{\infty} x^4 d\, F_o(x;\theta^o)$ exists, nor is it required that θ_N^* be given by (2.1). iii) An immediate consequence of assumption 2 is that $F_o^{-1}(t)$ is continuously differentiable for $0 < t < 1$ and that $dF_o^{-1}(t)/dt$ is given by $1/f_o(F_o^{-1}(t))$. Thus, it follows from Taylor's theorem with Cauchy remainder that for $0 < t < 1$ and $0 < \delta < 1$,

(2.2) $$F_o^{-1}(t) = F_o^{-1}(\delta) + (t-\delta)/f_o(F_o^{-1}(\delta)) + r(t,\delta),$$

where

(2.3) $$\sup_{|t-\delta|\leq\tau} |r(t,\delta)| = o(\tau) \quad \text{as} \quad \tau \to o.$$

iv) It will be shown in the appendix that assumptions 1 and 2 are sufficient to ensure that the assumptions given in section 2 of Chibisov (op.cit.) are satisfied.

3. The Quadratic Form Minimized by $\tilde{\theta}_N$ and Some Prelimary Results

The Pearson chi-square test statistic measures deviations from H_o by considering the difference between the empirical distribution function, $F_N(x)$, and the (possibly estimated) hypothesized distribution function, $F_o(x;\theta^o)$, evaluated at k-1 points interior to the support of F_o. Here, the empirical distribution function is defined as

(3.1) $$F_N(x) = \frac{1}{N}\sum_{\alpha=1}^{N} I_{(-\infty,x]}(X_\alpha)$$

where $I_A(x)$ denotes the indicator function of the set A.

For the fixed interval case, these points are specified prior to observing the sample and will be denoted by

(3.2) $$-\infty < c_1 < c_2 < \cdots < c_{k-1} < +\infty,$$

where it will be assumed that $k > 4$. Note that if c_o and c_k are defined to be $-\infty$ and $+\infty$ respectively, then assumption 2 guarantees that $\Delta F_o(c_i,\theta) > 0$, $i = 1,\ldots,k$, for all θ in Θ.

Define the (k-1)-dimensional random vector $u(\theta)$ by

(3.3) $$u_i(\theta) = F_N(c_i) - F_o(c_i;\theta), \quad i = 1,\ldots,k-1,$$

and define the $(k-1) \times (k-1)$ symmetric matrix $D(\theta)$ by

(3.4) $$D(\theta) = \{F_o(c_{\min(i,j)};\theta)[1-F_o(c_{\max(i,j)};\theta)]\}_{i,j=1}^{k-1}.$$

Then the minimum chi-square estimate, $\tilde{\theta}_N$, is defined to be any estimate for which there exists a sequence of positive constants ρ_N, with $\rho_N \to 0$ as $N \to \infty$, such that

(3.5) $$X_N^2(\tilde{\theta}_N) \leq \inf_{\theta \in \Theta} X_N^2(\theta) + \rho_N$$

where

(3.6) $$X_N^2(\theta) = N \sum_{i=1}^{K} (\Delta F_N(c_i) - \Delta F_o(c_i;\theta))^2/\Delta F_o(c_i;\theta)$$

$$= Nu(\theta)'D(\theta)^{-1}u(\theta).$$

Let $B = (b_{i,j})$ denote the $k \times 2$ matrix with $b_{i,j}$ given by

(3.7) $$b_{i,j} = (\Delta F_o(c_i;\theta^o))^{-1/2} \frac{\partial}{\partial\theta_j} \bar{\Delta} F_o(c_i;\theta)\Big|_{\theta=\theta^o} \quad i = 1,\ldots,k_i;\ j = 1,2,$$

and let $Y(\theta)$ denote the k-dimensional random vector with

(3.8) $$Y_i(\theta) = \bar{\Delta}u_i(\theta)/[\Delta F_o(c_i;\theta)]^{1/2}, \quad i = 1,\ldots,k.$$

If $\tilde{\theta}_N$ satisfies (3.5) and if assumptions 1 and 2 hold then (Chibisov (op.cit.) Theorem 5.1)

(3.9) $$\tilde{\theta}_N - \theta^o = (B'B)^{-1}B'Y(\theta^o) + o_p(1/\sqrt{N}).$$

Notice that

(3.10) $$\frac{\partial}{\partial\theta_1} F_o(x;\theta) = f_o(\theta_1+\theta_2 x)$$
$$\frac{\partial}{\partial\theta_2} F_o(x;\theta) = xf_o(\theta_1+\theta_2 x).$$

Thus, if the $(k-1) \times (k-1)$ symmetric matrices $\Pi(\theta)$ and $\Gamma(\theta)$ are defined by

(3.11) $$\Pi(\theta) = \mathrm{diag}[f_o(\theta_1+\theta_2 c_1), \ldots, f_o(\theta_1+\theta_2 c_{k-1})]$$
$$\Gamma(\theta) = \Pi(\theta)^{-1}D(\theta)\Pi(\theta)^{-1},$$

then (3.9) can be written as

$$\tilde{\theta}_N - \theta^o = [(1,c)'\Gamma(\theta^o)^{-1}(1,c)]^{-1}(1,c)'\Gamma(\theta^o)^{-1}\Pi(\theta^o)^{-1}u(\theta^o) + o_p(1/\sqrt{N}), \tag{3.12}$$

where 1 is the $k-1$ dimensional vector of ones and $c = (c_1,\ldots,c_{k-1})'$. Therefore, it follows from (3.12) that $\tilde{\theta}_N$ satisfies condition A with

$$h(x) = [(1,c)'\Gamma(\theta^o)^{-1}(1,c)]^{-1}(1,c)'\Gamma(\theta^o)^{-1}\Pi(\theta^o)^{-1}[I_{(-\infty,c]}(x) - F_o(c;\theta^o)] \tag{3.13}$$

where, for any vector $a = (a_1,\ldots,a_{k-1})'$,

$$I_{(-\infty,a]}(x) = (I_{(-\infty,a_1]}(x),\ldots,I_{(-\infty,a_{k-1}]}(x))' \tag{3.14}$$

and

$$F_o(a;\theta^o) = (F_o(a_1;\theta^o),\ldots,F_o(a_{k-1},\theta^o))'. \tag{3.15}$$

For the random interval case, as considered here, the points where the difference between F_N and F_o is observed are obtained as follows. First, prior to observing the sample, partition the interval $[0,1]$ into $k(> 4)$ intervals and let

$$0 = \delta_o^o < \delta_1^o \ldots < \delta_{k-1}^o < \delta_k^o = 1 \tag{3.16}$$

denote the end points of these intervals. Next, define the functions

$$g_i(\theta) = F_o^{-1}(\delta_i^o;\theta), \quad i = 1,\ldots,k-1. \tag{3.17}$$

Thus, for the parameterization specified by assumption 1, $g_i(\theta) = (\bar{c}_i-\theta_1)/\theta_2$ where $\bar{c}_i = F_o^{-1}(\delta_i^o)$, $i = 1,\ldots,k-1$. Then, after observing the sample, obtain a preliminary estimate, θ_N^*, of θ and compute $g_i(\theta_N^*)$ for $i = 1,\ldots,k-1$. The difference between F_N and F_o is then observed at the $k-1$ random points $g_i(\theta_N^*)$.

For θ and $\bar{\theta}$ in Θ, define the $(k-1)$-dimensional random vector $u(\bar{\theta},\theta)$ by

$$u_i(\bar{\theta},\theta) = F_N(g_i(\bar{\theta})) - F_o(g_i(\bar{\theta});\theta), \quad i = 1,\ldots,k-1, \tag{3.18}$$

and define the $(k-1) \times (k-1)$ symmetric matrix $D(\bar{\theta},\theta)$ by

$$D(\bar{\theta},\theta) = (F_o(g_{\min(i,j)}(\bar{\theta});\theta)\ [1-F_o(g_{\max(i,j)}(\bar{\theta});\theta)])_{i,j=1}^{k-1}. \tag{3.19}$$

Note that if $\theta = \bar{\theta}$ then

$$u_i(\theta,\theta) = F_N(g_i(\theta)) - \delta_i$$

$$D(\theta,\theta) = D = (\delta^o_{\min(i,j)}[1-\delta^o_{\max(i,j)}])_{i,j=1}^{k-1}. \tag{3.20}$$

Next, for θ and $\bar{\theta}$ in Θ, let $R_N^2(\bar{\theta},\theta)$ denote the quadratic form

$$(3.21) \qquad R_N^2(\bar{\theta},\theta) = N\, u(\bar{\theta},\theta)' D(\bar{\theta},\theta)^{-1} u(\bar{\theta},\theta).$$

If θ_N^* denotes the preliminary estimate of θ used to compute the random interval boundary points, then $\tilde{\theta}_N$ is defined to be any estimate for which there exists a sequence of positive constants ρ_n, where $\rho_N \to 0$ as $N \to \infty$, such that

$$(3.22) \qquad R_N^2(\theta_N^*,\tilde{\theta}_N) \leq \inf_{\theta \in \Theta} R_N^2(\theta_N^*,\theta) + \rho_N.$$

For θ and $\bar{\theta}$ in Θ define the $(k-1) \times (k-1)$ symmetric matrices $\Pi(\bar{\theta},\theta)$ and $\Gamma(\bar{\theta},\theta)$ by

$$(3.23) \qquad \begin{aligned} \Pi(\bar{\theta},\theta) &= \mathrm{diag}[f_o(\theta_1+\theta_2 g_1(\bar{\theta})),\ldots,f_o(\theta_1+\theta_2 g_{k-1}(\bar{\theta}))] \\ \Gamma(\bar{\theta},\theta) &= \Pi(\bar{\theta},\theta)^{-1} D(\bar{\theta},\theta)\Pi(\bar{\theta},\theta)^{-1} \end{aligned}$$

and notice that when $\theta = \bar{\theta}$

$$(3.24) \qquad \begin{aligned} \Pi(\theta,\theta) &= \Pi = \mathrm{diag}[f_o(\bar{c}_1),\ldots,f_o(\bar{c}_{k-1})] \\ \Gamma(\theta,\theta) &= \Gamma = \Pi^{-1} D \Pi^{-1}. \end{aligned}$$

Then, if $\theta_N^* - \theta^o = o_p(1)$, if assumptions 1 and 2 hold, and if $\tilde{\theta}_N$ satisfies (3.22) it can be shown (Chibisov (op.cit) Theorem 5.2) that

$$(3.25) \qquad \tilde{\theta}_N - \theta^o = [(1,g(\theta^o))'\Gamma^{-1}(1,g(\theta^o))]^{-1}(1,g(\theta^o))'\Gamma^{-1}\Pi^{-1}u(\theta^o,\theta^o) + o_p(1/\sqrt{N})$$

where

$$(3.26) \qquad g(\theta) = \frac{1}{\theta_2}[\bar{c}-\theta_1 1]$$

and $\bar{c} = (\bar{c}_1,\ldots,\bar{c}_{k-1})'$. Thus, it follows from (3.25) that $\tilde{\theta}_N$ satisfies condition A with $h(x)$ given by

$$(3.27) \qquad h(x) = [(1,g(\theta^o))'\Gamma^{-1}(1,g(\theta^o))]^{-1}(1,g(\theta^o))'\Gamma^{-1}\Pi^{-1}[I_{(-\infty,g(\theta^o)]}(x) - \delta^o],$$

where $\delta^o = (\delta_1^o,\ldots,\delta_{k-1}^o)'$.

If θ_N^* and θ_N^{**} denote two (possibly distinct) estimates of θ, then the Pearson chi-square test statistic is given by $X_N^2(\theta_N^{**})$ in the fixed interval case and by $R_N^2(\theta_N^*,\theta_N^{**})$ in the random interval case. In order to determine the limit distributions of these statistics under H_o, it is necessary to investigate the limit distributions of $u(\theta_N^{**})$ and $u(\theta_N^*,\theta_N^{**})$. If $\theta_N^* - \theta^o = o_p(1)$ and θ_N^{**} satisfies condition A then

III follows from assumptions 1 and 2 note that by Taylor's theorem with Cauchy remainder there exists an α, with $0 < \alpha < 1$, such that if $\bar{\theta} = \theta^o + \alpha(\theta-\theta^o)$ then

$$F_o(c_i;\theta) = F_o(c_i;\theta^o) + (\theta-\theta^o)'(1,c_i)'f_o((1,c_i)\theta^o) + r_i(\theta)$$

where $r_i(\theta) = (\theta-\theta^o)'(1,c_i)[f_o((1,c_i)\bar{\theta}) - f_o((1,c_i)\theta^o)]$. Thus, it follows from the Cauchy-Schwarz inequality that $|r_i(\theta)|/\|\theta-\theta^o\| \le \sqrt{1+c_i^2}\,|f_o((1,c_i)\theta) - f_o((1,c_i)\theta^o)|$. Therefore, since $\bar{\theta} \to \theta^o$ as $\bar{\theta} \to \theta^o$ as $\|\theta-\theta^o\| \to 0$ and since f_o is continuous, III is clear.

In the random interval case specify

(A.6) $$o = \delta_o^o < \delta_1^o < \dots < \delta_{k-1}^o < \delta_k^o = 1,$$

and for $\varepsilon > 0$ let $B(\varepsilon) = \cup_{i=1}^{k-1}(\delta_i^o-\varepsilon,\delta_i^o+\varepsilon)$ and define $s(\varepsilon)$ as in (A.5). Then Chibisov assumes

I) θ^o is in the interior of Θ.

II) $F_o(x;\theta^o)$ is continuous in x for $x \in S(\varepsilon)$.

III) There exists $\varepsilon > 0$ such that for k in $S(\varepsilon)$, $F_o(x;\theta)$ is differentiable in θ at θ^o and

$$F_o(x;\theta) = F_o(x;\theta^o) + (\theta-\theta^0)'h(x) + r(x,\theta)$$

where $h(x) = (h(x),\dots,h_s(x))'$, $h_j(x) = \frac{\partial}{\partial\theta_j} F_o(x;\theta)\big|_{\theta=\theta^0}$, $v_j(u) = h_j(F_o^{-1}(u;\theta^o))$ is bounded and continuous in u for u in $B(\varepsilon)$, and $\sup_{x\in S(\varepsilon)} |r(x,\theta)| = o(\|\theta-\theta^o\|)$ as $\|\theta-\theta^o\| \to 0$.

IV) The rank of $B = (\bar{\Delta}h_j(g_i(\theta^o))/\sqrt{\Delta\delta_i^o})_{i=1\ j=1}^{k\ \ s}$ is s.

V) For any $\tau > 0$ there exist $\rho > 0$ and $\varepsilon > 0$ such that if $x_i \in S(\delta_i^o,\varepsilon)$ for $i = 1,\dots,k-1$ and for $x_o = -\infty$, $x_k = +\infty$, then $\|\theta-\theta^o\| > \tau$ implies $\max_{1\le i\le k} |\Delta F_o(x_i;\theta^o) - \Delta\delta_i^o| > \rho$.

Clearly, assumptions I, II, IV, and V follow from assumptions 1 and 2 as in the fixed interval case. To show III the following notation will be used. Let

$\beta = \min(\delta_1^o, 1-\delta_{k-1}^o) > 0$ and for $0 < \varepsilon < \beta$ let

$D(\varepsilon) = \{x : (F_o^{-1}(\delta_1^o-\varepsilon)-\theta_1^o)/\theta_2^o \le x \le (F_o^{-1}(\delta_{k-1}^o+\varepsilon)-\theta_1^o)/\theta_2^o\}$ and define

$M(\varepsilon) = \max_{x\in D(\varepsilon)} \sqrt{1+x^2}$ which is finite since $D(\varepsilon)$ is compact. Also, for $\tau > 0$ and

APPENDIX

The purpose of this appendix is to show that the assumptions given in section 2 of Chibisov (op.cit.) (which are required to show (3.9) and (3.25)) are implied by assumptions 1 and 2. In order to accomplish this first note that for all θ in Θ, assumptions 1 and 2 imply that

(A.1) $$\frac{\partial}{\partial\theta} F_o(x;\theta) = (1,x)'f_o((1,x)\theta)$$

for all finite x and

$$F_o^{-1}(u;\theta) = (F_o^{-1}(u) - \theta_1)/\theta_2$$

for all $0 < u < 1$.

In the fixed interval case specify

(A.3) $$-\infty < c_1 < c_2 < \ldots < c_{k-1} < +\infty$$

and define

(A.4) $$\delta_i^o = F_o(c_i;\theta^o), \quad i = 1,\ldots,k-1.$$

In addition, for $\varepsilon > 0$ let $S(\varepsilon) = \cup_{i=1}^{k-1} S(\delta_i^o,\varepsilon)$ where

(A.5) $$S(\delta_i^o,\varepsilon) = \{x: |F_o(x,\theta^o) - \delta_i^o| < \varepsilon\}, \quad i = 1,\ldots,k-1.$$

Then Chibisov assumes:

I) θ^o is in the interior of Θ

II) $F_o(x;\theta^o)$ is continuous in x for $x \in S(\varepsilon)$

III) For each $i = 1,2,\ldots,k-1$, $F_o(c_i;\theta)$ is differentiable in θ at θ^o and

$$F_o(c_i;\theta) = F_o(c_i;\theta^o) + (\theta-\theta^o)'h(c_i) + r_i(\theta)$$

where $h(c_i) = (h_1(c_i),\ldots,h_s(c_i))'$, $h_j(c_i) = \frac{\partial}{\partial\theta_j} F_o(c_i;\theta)\big|_{\theta=\theta^0}$,

and $|r_i(\theta)| = o(\|\theta-\theta^o\|)$ as $\|\theta-\theta^o\| \to 0$.

IV) The rank of B is s where $B = \left(\bar{\Delta}h_j(c_i)/\sqrt{\Delta\delta_i^o}\right)_{i=1}^{k}{}_{j=1}^{s}$.

V) For any $\tau > 0$ there exists $\rho > 0$ such that

$\|\theta-\theta^o\| > \tau$ implies $\max_{1\le i\le k} |\Delta F_o(c_{ij}\theta) - \Delta\delta_i^o| > \rho$.

Clearly, assumptions 1 and 2 imply I and II since Θ is an open set and for all θ in Θ, $F_o(x;\theta)$ is continuous in x for all real x. In addition, (2.2) shows that V holds and since $k \geq 4$ the rank of B is 2 (see (3.7)) which shows IV. To see that

(7.5) $$\pi(\theta) = \mathrm{diag}[q_o(\theta_1+\theta_2 t(c_1)),\ldots,q_o(\theta_1+\theta_2 t(c_{k-1}))],$$

and, in (5.7), replace f_o with q_o, F_o with Q_o, and c_i with $t(c_i)$. Then, if $\theta_N^* - \theta^o = o_p(1)$, $\tilde{\theta}_N$ is given by (5.9) and the results of section 6 hold.

$$p(x) = \begin{cases} (\gamma_1/\gamma_2)(x/\gamma_2)^{\gamma_1 - 1} \exp\{-(x/\gamma_2)^{\gamma_1}\} & x > 0 \\ 0 & x < 0 \end{cases}$$

where γ_1 and γ_2 are positive, satisfies assumptions 3 and 4 with $A = (o,\infty)$, $Q_o(y) = 1-e^{-e^y}$, $\theta_1 = -\gamma_1 \log \gamma_2$, $\theta_2 = \gamma_1$, and $t(x) = \log(x)$.

The analysis under assumptions 3 and 4 is precisely the same as under assumptions 1 and 2 with the following modifications. In the random interval case i) redefine $\bar{c}_i = F_o^{-1}(\delta_i^o)$ and $g_i(\theta) = (\bar{c}_i - \theta_1)/\theta_2$ as

$$\bar{c}_i = Q_o^{-1}(\delta_i^o)$$

$$(7.1) \qquad v_i(\theta) = (\bar{c}_i - \theta_1)/\theta_2, \qquad i = 1,\dots,k-1$$

$$g_i(\theta) = t^{-1}(v_i(\theta)),$$

ii) change the definition of $\pi(\bar{\theta},\theta)$ in (3.23) to

$$(7.2) \qquad \pi(\bar{\theta},\theta) = \mathrm{diag}[q_o(\theta_1+\theta_2 v_1(\bar{\theta})),\dots,q_o(\theta_1+\theta_2 v_{k-1}(\bar{\theta}))],$$

iii) change (4.4) and (4.5) to

$$\bar{y}_i = \bar{c}_i + (1/q_o(\bar{c}_i))[F_N(g_i(\theta_N^*)) - \delta_i^o], \qquad i = 1,\dots,k-1$$

(7.3)

$$\bar{\varepsilon}(\theta) = \bar{y} - [1,v(\theta_N^*)]\theta,$$

where $v(\theta) = (v,(\theta),\dots,v_{k-1}(\theta))'$, and iv) in (4.10) and (4.11) replace f_o by q_o. If $\theta_N^* - \theta^o = o_p(1/\sqrt{N})$ then $\bar{\bar{\theta}}_N$ is given by (4.13) and the results of section 6 hold. In the fixed interval case replace (5.3) and (5.4) with

$$y_i = (1,t(c_i))\theta_N^* + (q_o((1,t(c_i))\theta_N^*))^{-1}[F_N(c_i) - F_o(c_i;\theta_N^*)], \quad i = 1,\dots,k-1$$

(7.4)

$$\varepsilon(\theta) = y - [1,t(c)]\theta$$

where $t(c) = (t(c_1),\dots,t(c_{k-1}))'$ and $\theta_N^* - \theta^o = o_p(1/\sqrt{N})$, change the definition of $\pi(\theta)$ given in (3.11) to

lemma 1, ii) the elements of $g(\theta)$ are continuous in θ, and iii)

(6.10) $$\hat{R}^2_N(\theta^*_N,\theta) - \hat{R}^2_N(\theta^*_N,\bar{\theta}) = N[(\bar{\theta}-\theta)'(1,g(\theta^*_N))'\Gamma^{-1}(\bar{\varepsilon}(\bar{\theta}) + (1,g(\theta^*_N))(\bar{\theta}-\theta))],$$

(6.9) shows that

(6.11) $$\hat{R}^2_N(\theta^*_N,\bar{\bar{\theta}}_N) - \hat{R}^2_N(\theta^*_N,\tilde{\theta}_N) = o_p(1).$$

Therefore, it follows from (6.3) and (3.27) that if $\theta^*_N - \theta^o = O_p(1/\sqrt{N})$ and if $\bar{\bar{\theta}}_N$ is given by (4.13) then

(6.12) $$R^2_N(\theta^*_N,\bar{\bar{\theta}}_N) - R^2_N(\theta^*_N,\tilde{\theta}_N) = o_p(1),$$

for any estimate $\tilde{\theta}_N$ satisfying (3.22).

Similary, if $\theta^*_N - \theta^o = o_p(1)$, if $\bar{\bar{\theta}}_N$ is given by (5.6), and if $\tilde{\theta}_N$ is any estimate satisfying (3.5), then, since i) the elements of $\Gamma(\theta)^{-1}$ are continuous in θ, ii) $\varepsilon(\bar{\bar{\theta}}_N) = O_p(1/\sqrt{N})$ by lemma 2 and (3.13), and iii)

(6.13) $$\hat{X}^2_N(\theta) - \hat{X}^2_N(\bar{\theta}) = N[(\bar{\theta}-\theta)'(1,c)'\Gamma(\theta^*_N)^{-1}(\varepsilon(\bar{\theta}) + (1,c)(\bar{\theta}-\theta))],$$

it follows from lemma 5, (3.13) and (6.5) that

(6.14) $$\bar{\bar{\theta}}_N - \tilde{\theta}_N = o_p(1/\sqrt{N}),$$

(6.15) $$\hat{X}^2_N(\bar{\bar{\theta}}_N) - \hat{X}^2_N(\tilde{\theta}_N) = o_p(1),$$

and

(6.16) $$X^2_N(\bar{\bar{\theta}}_N) - X^2_N(\tilde{\theta}_N) = o_p(1).$$

7. An Extension

In this section assumptions 1 and 2 are modified as follows.

Assumption 3: There exists a parameterization of F_o such that

i) $\Theta = \{\theta : \theta_1 \in \mathbb{R}$ and $\theta_2 > 0\}$

ii) The support of $F_o(\cdot;\theta)$ is the connected set A for all θ in Θ

iii) $F_o(x;\theta) = Q_o(\theta_1+\theta_2 t(x))$ where t is a homeomorphism from A onto $\mathbb{R}$.

Assumption 4: $Q_o(y)$ is a one to one continuously differentiable function of y and $q_o(y) = \frac{d}{dy} Q_o(y)$ is non-zero for all finite y.

As an example, the Weibull distribution, which has the density

(6.7) $$X_N^2(\tilde{\theta}_N) - \hat{X}_N^2(\tilde{\theta}_N) = o_p(1)$$

and

(6.8) $$X_N^2(\tilde{\theta}_N) \xrightarrow{d} \chi_{k-3}^2.$$

Lemma 5: i) If assumptions 1 and 2 hold, if $\theta_N^* - \theta^o = O_p(1/\sqrt{N})$, and if $\tilde{\theta}_N$ is given by (4.13) then $\tilde{\theta}_N$ satisfies condition A with $h(x)$ given by (3.27).

ii) If assumptions 1 and 2 hold, if $\theta_N^* - \theta^o = o_p(1)$, and if $\tilde{\theta}_N$ is given by (5.9) then $\tilde{\theta}_N$ satisfies condition A with $h(x)$ given by (3.13).

Proof: i) (4.5) and (4.8) imply that

$$\tilde{\theta}_N - \theta^o = [(1,g(\theta_N^*))'\Gamma^{-1}(1,g(\theta_N^*))]^{-1}(1,g(\theta_N^*))'\Gamma^{-1}\bar{\varepsilon}(\theta^o).$$

Thus, since the elements of $g(\theta)$ are continuous in θ and since $\bar{\varepsilon}(\theta^o) = O_p(1/\sqrt{N})$ if $\theta_N^* - \theta^o = O_p(1/\sqrt{N})$,

$$\tilde{\theta}_N - \theta^o = [(1,g(\theta^o))'\Gamma^{-1}(1,g(\theta^o))]^{-1}(1,g(\theta^o))'\Gamma^{-1}\bar{\varepsilon}(\theta^o) + o_p(1/\sqrt{N}).$$

Furthermore, $\bar{\varepsilon}(\theta^o) = \Pi^{-1}u(\theta_N^*,\theta^o) + o_p(1/\sqrt{N})$ by lemma 1 and (see Chibisov (op. cit.) corollary 3.3) $u(\theta_N^*,\theta^o) - u(\theta^o,\theta^o) = o_p(1/\sqrt{N})$. Thus,

$$\tilde{\theta}_N - \theta^o = [(1,g(\theta^o))'\Gamma^{-1}(1,g(\theta^o))]^{-1}(1,g(\theta^o))'\Gamma^{-1}\Pi^{-1}u(\theta^o,\theta^o) + o_p(1/\sqrt{N})$$

which is the same as (3.25).

ii) (5.4) and (5.6) imply that

$$\begin{aligned}\tilde{\theta}_N - \theta^o &= [(1,c)'\Gamma(\theta_N^*)^{-1}(1,c)]^{-1}(1,c)'\Gamma(\theta_N^*)^{-1}\varepsilon(\theta^o)\\ &= [(1,c)'\Gamma(\theta^o)^{-1}(1,c)]^{-1}(1,c)'\Gamma(\theta^o)^{-1}\varepsilon(\theta^o) + o_p(1/\sqrt{N})\end{aligned}$$

since i) $\theta_N^* - \theta^o = o_p(1)$, ii) the elements of $\Gamma(\theta)^{-1}$ are continuous in θ, and iii) $\varepsilon(\theta^o) = O_p(1/\sqrt{N})$ by lemma 2. Furthermore, (i) of lemma 2 shows that $\varepsilon(\theta^o) = \Pi(\theta^o)^{-1}u(\theta^o) + o_p(1/\sqrt{N})$. Thus,

$$\tilde{\theta}_N - \theta^o = [(1,c)'\Gamma(\theta^o)^{-1}(1,c)]^{-1}(1,c)'\Gamma(\theta^o)^{-1}\Pi(\theta^o)^{-1}u(\theta^0) + o_p(1/\sqrt{N})$$

which is the same as (3.12). ∎

An immediate consequence of lemma 5 is that if $\theta_N^* - \theta^o = O_p(1/\sqrt{N})$ and if $\tilde{\theta}_N$ is given by (4.13), then

(6.9) $$\tilde{\theta}_N - \hat{\theta}_N = o_p(1/\sqrt{N})$$

for any estimate $\hat{\theta}_N$ satisfying (3.22). Thus, since i) $\bar{\varepsilon}(\hat{\theta}_N) = O_p(1/\sqrt{N})$ by (3.27) and

Since the elements of $\Gamma(\theta)^{-1}$ are continuous in θ, $P(\theta_N^*) - P(\theta^o) = o_p(1)$. Thus, since $\varepsilon(\theta^o) = y - (1,c)\theta^o$ and since $P(\theta_N^*)(1,c) = 0$, $\varepsilon(\tilde{\theta}_N) = P(\theta_N^*)y = P(\theta_N^*)\varepsilon(\theta^o)$. Thus, $\hat{X}_N^2(\tilde{\theta}_N) = N\,\varepsilon(\theta^o)'P(\theta_N^*)'\Gamma(\theta_N^*)^{-1}P(\theta_N^*)\varepsilon(\theta^o)$ so that $\hat{X}_N^2(\tilde{\theta}_N) - N\varepsilon(\theta^o)'P(\theta^o)'\Gamma(\theta^o)^{-1}P(\theta^o)\varepsilon(\theta^o) = o_p(1)$ by lemma 2. Therefore, since $\sqrt{N}\,\varepsilon(\theta^o) \xrightarrow{d} N_{k-1}(o,\Gamma(\theta^o))$ and since $\Gamma(\theta^o)P(\theta^o)'\Gamma(\theta^o)^{-1}P(\theta^o)$ is idempotent with rank k-3, the result follows. ∎

In order to derive the limit distribution of $R_N^2(\theta_N^*,\tilde{\theta}_N)$ notice that (i) of lemma 1, (3.21), and (4.7) show that

(6.1) $$R_N^2(\theta_N^*,\theta) - \hat{R}_N^2(\theta_N^*,\theta) = \zeta(\theta_N^*,\theta)$$

where

(6.2) $$\zeta(\theta_N^*,\theta) = N\,\bar{\varepsilon}(\theta)'[\Gamma(\theta_N^*,\theta)^{-1} - \Gamma^{-1}]\,\bar{\varepsilon}(\theta) + N[2\bar{\varepsilon}(\theta) + e(\theta_N^*,\theta)]'\Gamma(\theta_N^*,\theta)^{-1}e(\theta_N^*,\theta).$$

Furthermore, if θ_N^{**} is an estimate of θ which satisfies condition A and if $\theta_N^* - \theta^o = O_p(1/\sqrt{N})$, lemma 1 and the continuity of the elements of $\Gamma(\theta_N^*,\theta)^{-1}$ show that $\zeta(\theta_N^*,\theta_N^{**}) = o_p(1)$. Thus, if it can be shown that $\tilde{\theta}_N$, as defined in (4.13), satisfies condition A, then

(6.3) $$R_N^2(\theta_N^*,\tilde{\theta}_N) - \hat{R}_N^2(\theta_N^*,\tilde{\theta}_N) = o_p(1)$$

and

(6.4) $$R_N^2(\theta_N^*,\tilde{\theta}_N) \xrightarrow{d} \chi_{k-3}^2.$$

Similarly,

(6.5) $$X_N^2(\theta) - \hat{X}_N^2(\theta) = \zeta(\theta)$$

where

(6.6) $$\zeta(\theta) = N\varepsilon(\theta)'[\Gamma(\theta)^{-1} - \Gamma(\theta_N^*)^{-1}]\varepsilon(\theta) + N[2\varepsilon(\theta) + r(\theta)]'\Gamma(\theta)^{-1}r(\theta)$$

and, if $\theta_N^* - \theta^o = o_p(1)$ and θ_N^{**} satisfies condition A, it follows from lemma 2 and the continuity of the elements of $\Gamma(\theta)^{-1}$ that $\zeta(\theta_N^{**}) = o_p(1)$. Thus, if $\tilde{\theta}_N$, as defined in (5.9), satisfies condition A, then

carrying out the arithmetic in (5.6), it is seen that

$$\tilde{\theta}_{1,N} = \hat{\alpha}(\theta_N^*) + \hat{\beta}(\theta_N^*)$$

(5.9)

$$\tilde{\theta}_{2,N} = \hat{\beta}(\theta_N^*)$$

where θ_N^* is a preliminary estimate of θ satisfying $\theta_N^* - \theta^o = o_p(1)$.

6. The Asymptotic Properties of $\tilde{\theta}_N$, $R_N^2(\theta_N^*,\tilde{\theta}_N)$, and $X_N^2(\tilde{\theta}_N)$.

The following two lemmas provide the limit distributions of $\hat{R}_N^2(\theta_N^*,\tilde{\theta}_N)$ and $\hat{X}_N^2(\tilde{\theta}_N)$ under H_o.

Lemma 3: If assumptions 1 and 2 hold, if $\theta_N^* - \theta^o = O_p(1/\sqrt{N})$, and if $\tilde{\theta}_N$ is given by (4.13) then under H_o,

$$\hat{R}_N^2(\theta_N^*,\tilde{\theta}_N) \xrightarrow{d} \chi^2_{k-3}.$$

Proof: Note first that the (i,j)th element of the matrix

$$[1,g(\theta)][(1,g(\theta))'\Gamma^{-1}(1,g(\theta))]^{-1}(1,g(\theta))'$$

is given by

$$(w_3 - (\bar{c}_i+\bar{c}_j)w_2 + \bar{c}_i\bar{c}_j w_3)/(w_1w_3 - w_2^2).$$

Thus, the $(k-1) \times (k-1)$ matrix

$$P = I - [1,g(\theta_N^*)]\left[(1,g(\theta_N^*))'\Gamma^{-1}(1,g(\theta_N^*))\right]^{-1}(1,g(\theta_N^*))'\Gamma^{-1}$$

does not depend on θ_N^*. Furthermore, since $\bar{\varepsilon}(\theta^o) = \bar{y} - [1,g(\theta_N^*)]\theta^o$ and since $P[1,g(\theta_N^*)] = 0$, $\bar{\varepsilon}(\tilde{\theta}_N) = P\bar{y} = P\,\bar{\varepsilon}(\theta^o)$. Thus, it follows from lemma 1 that $\sqrt{N}\,\bar{\varepsilon}(\tilde{\theta}_N) \xrightarrow{d} N_{k-1}(o,P\Gamma P')$ and, since $(P\Gamma P')\Gamma^{-1}$ is idempotent with rank $k-3$, the result follows. ∎

Lemma 4: If assumptions 1 and 2 hold, if $\theta_N^* - \theta^o = o_p(1)$, and if $\tilde{\theta}_N$ is given by (5.9) then under H_o,

$$\hat{X}_N^2(\tilde{\theta}_N) \xrightarrow{d} \chi^2_{k-3}.$$

Proof: Define the $(k-1) \times (k-1)$ matrix $P(\theta)$ by

$$P(\theta) = I - [1,c]\left[(1,c)'\Gamma(\theta)^{-1}(1,c)\right]^{-1}(1,c)'\Gamma(\theta)^{-1}.$$

Furthermore, since $\theta_N^{**} - \theta^o = O_p(1/\sqrt{N})$, so that $r(F_o(c_i;\theta_N^{**}),F_o(c_i;\theta^o)) = o_p(1/\sqrt{N})$, $r(\theta_N^{**}) = o_p(1/\sqrt{N})$ and (i) is proved. (ii) and (iii) now follow immediately from (i) and (3.28). ∎

If $\Gamma(\theta^o)$ were known then, in analogy with (4.7), $\breve{\theta}_N$ would be defined as that estimate of θ^o which minimizes the quadratic form $N\varepsilon(\theta)'\Gamma(\theta^o)^{-1}\varepsilon(\theta)$. Since $\Gamma(\theta^o)^{-1}$ is not known, however, it is necessary to estimate it. Thus, let θ_N^* denote the preliminary estimate of θ^o used to define y, and define

$$\hat{X}_N^2(\theta) = N\varepsilon(\theta)'\Gamma(\theta_N^*)^{-1}\varepsilon(\theta). \tag{5.5}$$

Then $\breve{\theta}_N$ minimizes $\hat{X}_N^2(\theta)$ with respect to θ and is given by

$$\breve{\theta}_N = [(1,c)'\Gamma(\theta_N^*)^{-1}(1,c)]^{-1}(1,c)'\Gamma(\theta_N^*)^{-1}y. \tag{5.6}$$

In analogy with section 4, let

$$\begin{aligned}
w_1(\theta) &= 1'\Gamma(\theta)^{-1}1 = \sum_{i=1}^{k} [\bar{\Delta}f_o(\theta_1+\theta_2c_i)]^2/\Delta F_o(\theta_1+\theta_2c_i \\
w_2(\theta) &= 1'\Gamma(\theta)^{-1}c = \sum_{i=1}^{k} [\bar{\Delta}f_o(\theta_1+\theta_2c_i)][\bar{\Delta}(c_if_o(\theta_1+\theta_2c_i))]/\Delta F_o(\theta_1+\theta_2c_i) \\
w_3(\theta) &= c'\Gamma(\theta)^{-1}c = \sum_{i=1}^{k} [\bar{\Delta}(c_if_o(\theta_1+\theta_2c_i))]^2/\Delta F_o(\theta_1+\theta_2c_i) \\
\hat{w}_2(\theta) &= 1'\Gamma(\theta)^{-1}y = \sum_{i=1}^{k} [\bar{\Delta}f_o(\theta_1+\theta_2c_i)][\bar{\Delta}(c_if_o(\theta_1+\theta_2c_i))]/\Delta F_o(\theta_1+\theta_2c_i) \\
\hat{w}_3(\theta) &= c'\Gamma(\theta)^{-1}y = \sum_{i=1}^{k} [\bar{\Delta}(c_if_o(\theta_1+\theta_2c_i))][\bar{\Delta}(y_if_o(\theta_1+\theta_2c_i))]/\Delta F_o(\theta_1+\theta_2c_i),
\end{aligned} \tag{5.7}$$

and define

$$\begin{aligned}
\hat{\alpha}(\theta) &= [w_3(\theta)\hat{w}_2(\theta) - w_2(\theta)\hat{w}_3(\theta)]/[w_1(\theta)w_3(\theta) - w_2(\theta)^2] \\
\hat{\beta}(\theta) &= [w_1(\theta)\hat{w}_3(\theta) - w_2(\theta)\hat{w}_2(\theta)]/[w_1(\theta)w_3(\theta) - w_2(\theta)^2].
\end{aligned} \tag{5.8}$$

Notice that y_i is a consistent estimate of $\theta_1^o + \theta_2^o c_i$, not c_i, so that $\hat{\alpha}(\theta)$ and $\hat{\beta}(\theta)$ do not have the same interpretations as in section 4. In fact, by

where $\pi_i(\theta^o) = f_o(\theta_1^o + \theta_2^o c_i)$. However, since θ^o is unknown, in order to define the fixed interval analog to the random vector $\bar{y}$ it is first necessary to estimate θ^o. Thus, let θ_N^* denote a preliminary estimate of θ^o satisfying $\theta_N^* - \theta^o = O_p(1/\sqrt{N})$, let $y = (y_1,\ldots,y_{k-1})'$ where

$$y_i = (1,c_i)\theta_N^* + \pi_i(\theta_N^*)^{-1}[F_N(c_i) - F_o(c_i;\theta_N^*)], \quad i = 1,\ldots,k-1, \tag{5.3}$$

and, for θ in Θ, define the (k-1)-dimensional random vector $\varepsilon(\theta)$ by

$$\varepsilon(\theta) = y - [1,c]\theta. \tag{5.4}$$

Then the following lemma records some properties of $\varepsilon(\theta)$.

Lemma 2: If $\theta_N^* - \theta^o = O_p(1/\sqrt{N})$, if θ_N^{**} satisfies condition A, and if assumptions 1 and 2 are satisfied, then

i) $\varepsilon(\theta) = \Pi(\theta^o)^{-1}u(\theta) + r(\theta)$ where $r(\theta_N^{**}) = o_p(1/\sqrt{N})$

ii) $\sqrt{N}\,\varepsilon(\theta_N^{**}) \xrightarrow{d} N_{k-1}(o,\Pi(\theta^o)^{-1}\ddagger_o(\theta^o)\Pi(\theta^o)^{-1})$ where $\ddagger_o(\theta^o)$ is defined in (3.28), and

iii) $\sqrt{N}\,\varepsilon(\theta^o) \xrightarrow{d} N_{k-1}(o,\Gamma(\theta^o))$.

Proof: Since $F_o(c_i;\theta)$ is continuously differentiable in θ with $\frac{\partial}{\partial\theta}F_o(c_i;\theta) = (1,c_i)'f_o(\theta_1+\theta_2 c_i)$,

$$\begin{aligned}
y_i &= (1,c_i)\theta_N^* + \pi_i(\theta_N^*)^{-1}[F_N(c_i) - F_o(c_i;\theta^o) - \{F_o(c_i;\theta_N^*) - F_o(c_i;\theta^o)\}] \\
&= (1,c_i)\theta_N^* + \pi_i(\theta_N^*)^{-1}[F_N(c_i) - F_o(c_i;\theta^o) - \pi_i(\theta^o)(1,c_i)(\theta_N^*-\theta^o) + o_p(1/\sqrt{N})] \\
&= (1,c_i)\theta^o + \pi_i(\theta^o)^{-1}[F_N(c_i) - F_o(c_i;\theta^o)] + o_p(1/\sqrt{N}).
\end{aligned}$$

Note that the last equality follows from the fact that $\pi_i(\theta)$ is continuous in θ and $\theta_N^* - \theta^o = o_p(1)$. Thus, it follows from (2.2) and (5.4) that

$$\begin{aligned}
\varepsilon_i(\theta) &= \pi_i(\theta^o)^{-1}[F_N(c_i) - F_o(c_i;\theta^o)]-[F_o^{-1}(F_o(c_i;\theta))-F_o^{-1}(F_o(c_i;\theta^o))] + o_p(1/\sqrt{N}) \\
&= \pi_i(\theta^o)^{-1}[F_N(c_i)-F_o(c_i;\theta^o)]-\pi_i(\theta^o)^{-1}[F_o(c_i;\theta)-F_o(c_i;\theta^o)]-r(F_o(c_i;\theta);F_o(c_i;\theta^o)) \\
&\qquad + o_p(1/\sqrt{N}) \\
&= \pi_i(\theta^o)^{-1}u_i(\theta) + r(\theta) \quad \text{where} \quad r(\theta) = -r(F_o(c_i;\theta),F_o(c_i;\theta^o)) + o_p(1/\sqrt{N}).
\end{aligned}$$

which can be computed prior to observing the sample. Next, since $\bar{y} - \bar{c} = O_p(1/\sqrt{N})$, define

$$\hat{w}_2 = 1'\Gamma^{-1}\bar{y} = \sum_{i=1}^{k} [\bar{\Delta}f_o(\bar{c}_i)][\bar{\Delta}(\bar{y}_i f_o(\bar{c}_i))]/\Delta\delta_i^o \tag{4.11}$$

$$\hat{w}_3 = \bar{c}'\Gamma^{-1}\bar{y} = \sum_{i=1}^{k} [\bar{\Delta}(\bar{y}_i f_o(\bar{c}_i))][\bar{\Delta}(\bar{c}_i f_o(\bar{c}_i))]/\Delta\delta_i^o.$$

Finally, let

$$\hat{\alpha} = (w_3\hat{w}_2 - w_2\hat{w}_3)/(w_1w_3 - w_2^2) \tag{4.12}$$

$$\hat{\beta} = (w_1\hat{w}_3 - w_2\hat{w}_2)/(w_1w_3 - w_2^2)$$

and note that $\hat{\alpha}$ is a consistent estimate of 0 and $\hat{\beta}$ is a consistent estimate of 1. Then, by carrying out the arithmetic involved in (4.8), it follows that

$$\tilde{\theta}_{1,N} = \hat{\alpha} + \hat{\beta}\theta^*_{1,N} \tag{4.13}$$

$$\tilde{\theta}_{2,N} = \hat{\beta}\theta^*_{2,N}\ .$$

Thus, regardless of the initial estimate, θ^*_N, employed (provided only that $\theta^*_N - \theta^o = O_p(1/\sqrt{N})$), $\tilde{\theta}_N$ is obtained from it by scaling with a consistent estimate of 1 and adding a consistent estimate of 0.

5. The Least Squares Estimate $\tilde{\theta}_N$: The Fixed Interval Case

Since the arguments leading to the definition of $\tilde{\theta}_N$ in the fixed interval case are similar to those used in the random interval case, the details will only be sketched. In the fixed interval case an estimate $\tilde{\theta}_N$ satisfying (3.5) can be found, under suitable regularity conditions, as the solution of

$$(1,c)'\Pi(\theta)D(\theta)^{-1}u(\theta) = 0. \tag{5.1}$$

The least squares approximation to the solution of (5.1) is defined as follows. First note that since $F_N(c_i) - F_o(c_i;\theta^o) = O_p(1/\sqrt{N})$, (2.2) shows that

$$F_o^{-1}(F_n(c_i)) - F_o^{-1}(F_o(c_i;\theta^o)) = \pi_i(\theta^o)^{-1}[F_N(c_i) - F_o(c_i;\theta^o)] + o_p(1/\sqrt{N}), \tag{5.2}$$

$F_o(g_i(\theta);\theta^o) = \delta_i^o - f_o(\bar{c}_i)(1,g_i(\theta^o))(\theta-\theta^o) + v(\theta,\theta^o)$ where $\sup_{\|\theta-\theta^o\|<\tau} |v(\theta,\theta^o)| = o(\tau)$ as $\tau \to o$. Thus,

$$F_o(g_i(\theta_N^*);\theta^o) - \delta_i^o = O_p(1/\sqrt{N}) \tag{4.6}$$

by the standard arguments due to Mann and Wald (1943). Therefore, since $F_N(g_i(\theta_N^*)) - \delta_i^o = u_i(\theta_N^*,\theta^o) + F_o(g_i(\theta_N^*);\theta^o) - \delta_i^o = O_p(1/\sqrt{N})$ by (3.28) and (4.6), $F_o(g_i(\theta_N^*);\theta_N^{**}) - \delta_i^o = F_N(g_i(\theta_N^*)) - \delta_i^o - u_i(\theta_N^*,\theta_N^{**}) = O_p(1/\sqrt{N})$ so that $r(F_o(g_i(\theta_N^*);\theta_N^{**});\delta_i^o) = o_p(1/\sqrt{N})$. Thus, $e_i(\theta_N^*,\theta_N^{**}) = o_p(1/\sqrt{N})$ and (i) is proved. (ii) and (iii) now follow immediately from (3.28). ∎

Define the quadratic form $\hat{R}_N^2(\theta_N^*,\theta)$ by

$$\hat{R}_N^2(\theta_N^*,\theta) = N\,\bar{\varepsilon}(\theta)'\Gamma^{-1}\bar{\varepsilon}(\theta). \tag{4.7}$$

Then the least squares approximation to $\tilde{\theta}_N$, $\ddot{\theta}_N$, minimizes $\hat{R}_N^2(\theta_N^*,\theta)$ with respect to θ. Since the asymptotic covariance matrix of $\sqrt{N}\,\bar{\varepsilon}(\theta^o)$ is Γ, it follows from the theory of the general linear model that $\ddot{\theta}_N$ is given by

$$\ddot{\theta}_N = [(1,g(\theta_N^*))'\Gamma^{-1}(1,g(\theta_N^*))]^{-1}(1,g(\theta_N^*))'\Gamma^{-1}\bar{y}. \tag{4.8}$$

In order to simplify (4.8) first note that since the (i,j) the element of Γ^{-1} is given by

$$\Gamma_{i,j}^{-1} = \begin{cases} f_o(\bar{c}_i)^2[1/\Delta\delta_i^o + 1/\Delta\delta_{i+1}^o] & \text{if } |i-j| = 0 \\ -f_o(\bar{c}_i)f_o(\bar{c}_{i+1})/\Delta\delta_{i+1}^o & \text{if } |i-j| = 1, \quad i,j = 1,\dots,k-1, \\ o & \text{if } |i-j| > 1 \end{cases} \tag{4.9}$$

straightforward calculations show that

$$\begin{aligned} w_1 &= 1'\Gamma^{-1}1 = \sum_{i=1}^{k} [\bar{\Delta} f_o(\bar{c}_i)]^2/\Delta\delta_i^o \\ w_2 &= 1'\Gamma^{-1}\bar{c} = \sum_{i=1}^{k} [\bar{\Delta} f_o(\bar{c}_i)][\bar{\Delta}(\bar{c}_i f_o(\bar{c}_i))]/\Delta\delta_i^o \\ w_3 &= \bar{c}'\Gamma^{-1}\bar{c} = \sum_{i=1}^{k} [\bar{\Delta}(\bar{c}_i f_o(\bar{c}_i))]^2/\Delta\delta_i^o, \end{aligned} \tag{4.10}$$

The principal idea behind the least squares approximation to the solution of (4.1) is the fact that for all θ in Θ,

(4.2) $$F_o^{-1}(F_o(g_i(\theta_N^*);\theta)) = (1,g_i(\theta_N^*)\theta, \quad i = 1,\ldots,k-1.$$

Thus, it is tempting to regress the (k-1)-dimensional random vector, with ith component given by $F_o^{-1}(F_N(g_i(\theta_N^*)))$, onto the matrix $[1,g(\theta_N^*)]$ to obtain an estimate of θ^o. However, $|F_o^{-1}(F_N(g_i(\theta_N^*)))|$ is infinite if $F_N(g_i(\theta_N^*))$ is equal to zero or one, and this can happen with positive probability. But, since $F_N(g_i(\theta_N^*)) - \delta_i^o = O_p(1/\sqrt{N})$ if $\theta_N^* - \theta^o = O_p(1/\sqrt{N})$ (see the proof of lemma 1 below), (2.2) shows that

(4.3) $$F_o^{-1}(F_N(g_i(\theta_N^*))) - \bar{c}_i = (\pi_i)^{-1}(F_N(g_i(\theta_N^*)) - \delta_i^o) + o_p(1/\sqrt{N}), \quad i = 1,\ldots,k-1,$$

where $\pi_i = f_o(\bar{c}_i)$, and $(\pi_i)^{-1}(F_N(g_i(\theta_N^*)) - \delta_i^o)$ is finite regardless of whether $F_N(g_i(\theta_N^*))$ equals zero or one.

Define the (k-1)-dimensional random vector $\bar{y}$ by

(4.4) $$\bar{y}_i = \bar{c}_i + (\pi_i)^{-1}(F_N(g_i(\theta_N^*)) - \delta_i^o), \quad i = 1,\ldots,k-1,$$

where θ_N^* satisfies $\theta_N^* - \theta^o = O_p(1/\sqrt{N})$, and notice that $\bar{y}_i - \bar{c}_i = O_p(1/\sqrt{N})$. Next, for θ in Θ define the (k-1)-dimensional random vector $\bar{\varepsilon}(\theta)$ by

(4.5) $$\bar{\varepsilon}(\theta) = \bar{y} - [1,g(\theta_N^*)]\theta.$$

Some properties of $\bar{\varepsilon}(\theta)$ are recorded in the following lemma.

<u>Lemma 1</u>: If $\theta_N^* - \theta^o = O_p(1/\sqrt{N})$, if θ_N^{**} satisfies condition A, and if assumptions 1 and 2 are satisfied, then

i) $\bar{\varepsilon}(\theta) = \Pi^{-1}\, u(\theta_N^*,\theta) + e(\theta_N^*,\theta_N^{**}) = o_p(1/\sqrt{N})$,

ii) $\sqrt{N}\;\bar{\varepsilon}(\theta_N^{**}) \xrightarrow{d} N_{k-1}(0,\Pi^{-1}\ddagger_o\Pi^{-1})$ where $\ddagger_o$ is defined in (3.28), and

iii) $\sqrt{N}\;\bar{\varepsilon}(\theta^o) \xrightarrow{d} N_{k-1}(0,\Gamma)$.

Proof: Since $\pi_i^{-1} = 1/f_o(\bar{c}_i) = \frac{d}{dt}F_o(t)\Big|_{t=\delta_i^o}$, it follows from (2.2) that

$$\begin{aligned}\bar{\varepsilon}_i(\theta) &= (\pi_i)^{-1}[F_N(g_i(\theta_N^*)) - \delta_i^o] - [F_o^{-1}(F_o(g_i(\theta_N^*);\theta)) - F_o^{-1}(\delta_i^o)] \\ &= (\pi_i)^{-1}[F_N(g_i(\theta_N^*)) - \delta_i^o] - (\pi_i)^{-1}[F_o(g_i(\theta_N^*);\theta) - \delta_i^o] - r(F_o(g_i(\theta_N^*);\theta),\delta_i^o) \\ &= (\pi_i)^{-1}u_i(\theta_N^*,\theta) + e_i(\theta_N^*,\theta) \quad \text{where} \quad e_i(\theta_N^*,\theta) = -r(F_o(g_i(\theta_N^*);\theta),\delta_i^o).\end{aligned}$$

Next, it follows from assumptions 1 and 2 and Taylor's theorem with Cauchy remainder that

assumptions 1 and 2 are sufficient to show that (see, e.g., Moore and Spruill (op.cit.))

$$\sqrt{N}\, u(\theta_N^{**}) \xrightarrow{d} N_{k-1}(o, \Sigma_o(\theta^o))$$

(3.28)

$$\sqrt{N}\, u(\theta_N^{*}, \theta_N^{**}) \xrightarrow{d} N_{k-1}(o, \Sigma_o).$$

Here, the notation Σ_o and $\Sigma_o(\theta^o)$ for the asymptotic covariance matrices of $\sqrt{N}\, u(\theta_N^*, \theta_N^{**})$ and $\sqrt{N}\, u(\theta_N^{**})$, respectively, has been employed to stress the fact that under the parameterization imposed by assumption 1, if θ_N^* and θ_N^{**} are invariant with respect to the group of linear transformations with positive slope, then Σ_o does not depend on θ^o, while $\Sigma_o(\theta^o)$ does depend on θ^o regardless of whether θ_N^{**} is invariant or not. For the present purpose, the particular form of Σ_o and $\Sigma_o(\theta^o)$ is not important except when $\theta_N^{**} = \theta^o$, in which case $\Sigma_o = D$ and $\Sigma_o(\theta^o) = D(\theta^o)$.

It now follows from (3.28) that the limit distributions of $X^2(\theta^{**})$ and $R_N^2(\theta_N^*, \theta_N^{**})$ are the same as the distribution of

(3.29)
$$\sum_{i=1}^{k-1} \lambda_i z_i^2$$

where $z_1, \ldots, z_{k-1}$ are independent standard normal random variables and $\lambda_1, \ldots, \lambda_{k-1}$ are the characteristic roots of $D^{-1}\Sigma_o$ in the random interval case and $D(\theta^o)^{-1}\Sigma_o(\theta^o)$ in the fixed interval case. Furthermore, it is well known that if $\theta_N^{**} = \tilde{\theta}_N$ then $k-3$ of the λ_i are one while the remaining two λ_i are zero.

4. The Least Squares Estimate $\tilde{\theta}_N$: The Random Interval Case

In the random interval case an estimate $\tilde{\theta}_N$ satisfying (3.22) can be found, under suitable regularity conditions, by equating to zero the derivatives of $R_N^2(\theta_N^*, \theta)$ with respect to θ (while treating θ in $D(\theta_N^*, \theta)^{-1}$ as fixed) and solving for θ. This leads to the system of equations

(4.1)
$$(1, g(\theta_N^*))'\Pi(\theta_N^*, \theta) D(\theta_N^*, \theta)^{-1}\, u(\theta_N^*, \theta) = 0$$

which is generally difficult to solve for $\tilde{\theta}_N$.

$0 < \varepsilon < \beta$ let $B(\theta^o,\tau) = \{\theta : \|\theta-\theta^o\| < \tau\}$ and let $A(\varepsilon,\tau) = \{y = (1,x)\theta : \theta \in B(\theta^o,\tau),\ x \in D(\varepsilon)\}$. Finally, let $\bar{B}(\varepsilon)$ denote the closure of the set $B(\varepsilon)$ defined in the line following (A.6).

It follows from assumptions 1 and 2 that for any ε with $0 < \varepsilon < \beta$, $F_o(x;\theta)$ is continuously differentiable in θ at θ^o for any x in $D(\varepsilon)$. Thus, by Taylor's theorem with Cauchy remainder there exists an α $(0 < \alpha < 1)$, which depends on x and θ, such that for $\bar{\theta} = \theta^o + \alpha(\theta-\theta^o)$,

$$F_o(x;\theta) = F_o(x;\theta^o) + (\theta-\theta^o)'(1,x)'f_o((1,x)\theta^o) + r(x,\theta)$$

where $r(x,\theta) = (\theta-\theta^o)'(1,x)'[f_o((1,x)\bar{\theta}) - f_o((1,x)\theta^o)]$. Furthermore, since $0 < \varepsilon < \beta$ implies that $\bar{B}(\varepsilon)$ is contained in the interval $(0,1)$ and since, for such ε, $v(u) = (1,(F_o^{-1}(u)-\theta_1^o)/\theta_2^o)'f_o(F_o^{-1}(u))$ is continuous, and thus bounded, on $\bar{B}(\varepsilon)$, it follows that $v(u)$ is continuous and bounded on $B(\varepsilon)$.

Choose ε_1 and τ_1 so that $0 < \varepsilon_1 < \beta$ and $\tau_1 > 0$. By the Cauchy-Schwarz inequality

$$|r(x,\theta)|/\|\theta-\theta^o\| < \sqrt{1+x^2}\ |f_o((1,x)\bar{\theta})-f_o((1,x)\theta^o)|$$

for all (x,θ) in $D(\varepsilon_1) \times B(\theta^o,\tau_1)$. Furthermore, since f_o is continuous on $A(\varepsilon_1,\tau_1)$ and since $A(\varepsilon_1,\tau_1)$ is compact, f_o is uniformly continuous on $A(\varepsilon_1,\tau_1)$. Thus, given $\varepsilon_2 > 0$ (with $\varepsilon_2 < \varepsilon_1$) there exists $t_2 > 0$ (with $\tau_2 < \tau_1$) such that $|(1,x)\bar{\theta} - (1,x)\theta^o| < \tau_2$ implies $|f_o((1,x)\bar{\theta}) - f_o((1,x)\theta^o)| < \varepsilon_2/M(\varepsilon_1)$. Furthermore, since $|(1,x)\bar{\theta} - (1,x)\theta^o| = \alpha|(1,x)(\theta-\theta^o)| < \sqrt{1+x^2}\ \|\theta-\theta^o\| < M(\varepsilon_1)\|\theta-\theta^o\|$ for x in $D(\varepsilon_1)$, $\|\theta-\theta^o\| < \tau_2/M(\varepsilon_1)$ implies that $|f_o((1,x)\bar{\theta}) - f_o((1,x)\theta^o)| < \varepsilon_2/M(\varepsilon_1)$ for all x in $D(\varepsilon_1)$. Therefore, if $\|\theta-\theta^o\| < \tau_2/M(\varepsilon_1)$ then $\sup_{x \in S(\varepsilon_1)} |r(x,\theta)|/\|\theta-\theta^o\| < M(\varepsilon_1)\varepsilon_2/M(\varepsilon_1) = \varepsilon_2$ since $S(\varepsilon_1) \subset D(\varepsilon_1)$. This shows that $\sup_{x \in S(\varepsilon_1)} |r(x,\theta)| = o(\|\theta-\theta^o\|)$ as $\|\theta - \theta^o\| \to o$ and, therefore, III follows from assumptions 1 and 2.

REFERENCES

[1] Birch, M. W. (1964). A new proof of the Fisher-Pearson theorem. Ann. Math. Statist. 35, 817-824.

[2] Chernoff, H. and E. L. Lehmann (1954). The use of maximum likelihood estimates in X^2 test for goodness of fit. Ann. Math. Statist. 25, 579-586.

[3] Chibisov, D. M. (1971). Certain chi-square type tests for continuous distributions. Theor. Probability Appl. 16, 1-22.

[4] Cramér, H. (1946). Mathematical Methods of Statistics. Princeton Univ. Press.

[5] Dahiya, R. C. and J. Gurland (1972). Pearson chi-squared test of fit with random intervals. Biometrika 59, 147-153.

[6] Dahiya, R. C. and J. Gurland (1973). How many classes in the Pearson chi-square test? J. Amer. Statist. Assoc. 68, 707-712.

[7] Fisher, R. A. (1924). The conditions under which X^2 measures the discrepancy between observations and hypothesis. J. Roy. Statist. Soc. 87, 442-450.

[8] Mann, H. B. and A. Wald (1943). On stochastic limit and order relationships. Ann. Math. Statist. 14, 217-226.

[9] Moore, D. S. (1977). Generalized inverses, Wald's method, and the construction of chi-square tests of fit. J. Amer. Statist. Assoc. 72, 131-137.

[10] Moore, D. S. and M. C. Spruill (1975). Unified large sample theory of general chi-square statistics for tests of fit. Ann. Statist. 3, 599-616.

[11] Rao, C. R. (1958). Maximum likelihood estimation for the multinomial distribution with infinite number of cells. Sankhyā 20, 211-218.

[12] Rao, K. C. and D. S. Robson (1974). A chi-square statistic for goodness-of-fit tests within the exponential family. Commun. Statist. 3, 1139-1153.

[13] Roy, A. R. (1956). On X^2 statistics with variable intervals. Thesis, Stanford Univ.

[14] Spruill, M. C. (1976a). Cell selection in the Chernoff-Lehmann chi-square statistic. Ann. Statist. 4, 375-383.

[15] Spruill, M. C. (1976b). A comparison of chi-square goodness-of-fit tests based on approximate Bahadur slope. Ann. Statist. 4, 409-412.

[16] Watson, G. S. (1957). The X^2 goodness-of-fit test for normal distributions. Biometrika 44, 336-348.

[17] Watson, G. S. (1958). On chi-square goodness-of-fit tests for continuous distributions. J. Roy. Statist. Soc. Ser. B 20, 44-61.

[18] Watson, G. S. (1959). Some recent results in X^2 goodness-of-fit tests. Biometrics 15, 440-468.

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER 2792	2. GOVT ACCESSION NO. AD-A153 534
3. RECIPIENT'S CATALOG NUMBER	
4. TITLE (and Subtitle) LEAST-SQUARES APPROXIMATION TO MINIMUM CHI-SQUARE ESTIMATORS OF LOCATION AND SCALE PARAMETERS AND THEIR EFFECT ON THE PEARSON CHI-SQUARE TEST	5. TYPE OF REPORT & PERIOD COVERED Summary Report - no specific reporting period
	6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) A. E. Muhly and J. Gurland	8. CONTRACT OR GRANT NUMBER(s) DAAG29-80-C-0041
9. PERFORMING ORGANIZATION NAME AND ADDRESS Mathematics Research Center, University of Wisconsin, 610 Walnut Street, Madison, Wisconsin 53706	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS Work Unit Number 4 - Statistics and Probability
11. CONTROLLING OFFICE NAME AND ADDRESS U. S. Army Research Office, P.O. Box 12211, Research Triangle Park, North Carolina 27709	12. REPORT DATE February 1985
	13. NUMBER OF PAGES 26
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)	15. SECURITY CLASS. (of this report) UNCLASSIFIED
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

Accession For
NTIS GRA&I ☒
DTIC TAB ☐
Unannounced ☐
Justification

By
Distribution/
Availability Codes
Avail and/or
Dist Special
A-1

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Pearson chi-square test
test of fit
asymptotic distribution
least-squares approximation

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Application of the Pearson chi-square test to goodness of fit of a distribution often leads to serious difficulties, particularly in the formation of intervals (as in the case of a continuous distribution) and in the estimation of unknown parameters. Under suitable conditions and with appropriately constructed estimators of the parameters, the test statistic converges in

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

20. ABSTRACT - cont'd.

distribution to that of chi-square as the sample size increases. In the present paper, a comparatively simple least-squares approximation to the minimum chi-square estimator is developed which, when appropriately implemented, results in an asymptotic chi-square distribution of the test statistic. This estimator is developed for the cases of fixed and random intervals, and the role of the underlying assumptions is studied in detail.

END

FILMED

6-85

DTIC